VPS VIBRATIONS IN PHYSICAL SYSTEMS

DOI: 10.21008/j.0860-6897.2021.2.15
Vibrations in Physical Systems, 32(2):2021215, 2021

Impact of the Level of Noise and Echo on the Reaction Time of Listeners in the Perception of Logatoms

Stefan BRACHMAŃSKI[1], Andrzej DOBRUCKI[1]

Corresponding author: Stefan BRACHMAŃSKI, email: Stefan.brachmanski@pwr.edu.pl

[1] Wroclaw University of Science and Technology, Faculty of Electronics, Photonics and Microsystems, Department of Acoustics, Multimedia and Signal Processing, 50-370 Wrocław, Wybrzeże Wyspiańskiego 27

Abstract The article presents the results of research regarding the impact of the degree of distortion and noise of the logatom (nonsense word) on the listener's reaction time. The study aimed to determine the maximum reaction time of listeners, which will allow determining the time after which the logatom will be exposed in the speech quality assessment method with an alternative choice. The research was carried out with the participation of a group of ten students. A strong relationship between the results obtained and the concentration of the listeners was found, as well as the effect of fatigue, training, and the gender of the listener. The obtained results indicate that in the method with an alternative choice before the logatom emission should appear 1.1 s initial sequence, which will eliminate the situation when the listeners did not recognize the initial phoneme transmitted from the logatom.

Keywords: perception, listener's time reaction, speech quality

1. Introduction

Measurement of reaction time is an important issue in scientific research in many fields, such as psychology, psycholinguistics, acoustics, as well as in diagnostics, e.g. audiometry, psychology, and sports [1-8]. In speech perception research, analysis of reaction time is not a new issue. Research on speech perception shows that vowels are recognized better than consonants. This is most likely because the acoustic energy of vowels is about 6 dB greater than that of consonants [9]. Additionally, the consonants are shorter than the vowels. The smaller amplitude and short duration of consonants make their perception more difficult. In assessing the quality of speech signal transmission, the following are used: logatoms[1] (nonsense words), most of which begin with a consonant. If the listeners are presented with logatom beginning with a consonant, and without the initial sequence (noise, carrier phrase), then there is a high probability of its erroneous recognition.

The listener's reaction time is understood here as the time needed for a person to hear a speech signal, process it and make a decision as to its content.

The issue of the listener's reaction time is very important in the subjective methods of measuring speech intelligibility, based on the computer technique of presenting test samples, which are registered in the computer's memory in the form of isolated test units, e.g. words or logatoms. At this point, there is the problem of the duration of the initial phase occurring before the actual test signal appears (Fig. 1).

Too short initial sequence results in no reception or incorrect reception of the first phoneme in the given logatom. This problem plays a particularly important role in the assessment of speech intelligibility carried out with the alternative choice method "modified intelligibility test with forced-choice" (MIT-FC) [10]. In subjective measurements of speech intelligibility, the so-called a carrying phrase, eg "Now you will hear the next logatom. Please write it down". The carrier phrase stabilizes the listening conditions, especially for high levels of interfering noise. In the MIT-FC method, individual test units are generated by a computer without a carrier phrase. The lack of a carrier phrase, as is the case with the MIT-FC method, means that the listener's hearing organ is not prepared to receive the next logatom. This often results in incorrect

[1] 1 Logatom - a sequence of sounds in succession according to certain linguistic rules, devoid of any semantic content; in English nonsense word.

recognition of the first phoneme or even its omission. This problem occurs especially when speech intelligibility measurements are made in conditions of significant noise in the speech. It is important to determine the minimal duration of the start sequence related to the listeners' response time to the speech.

The aim of the presented research was:

- determining the reaction time of listeners necessary for correct logatom reception,
- examine the impact of disturbances and distortions on response time.

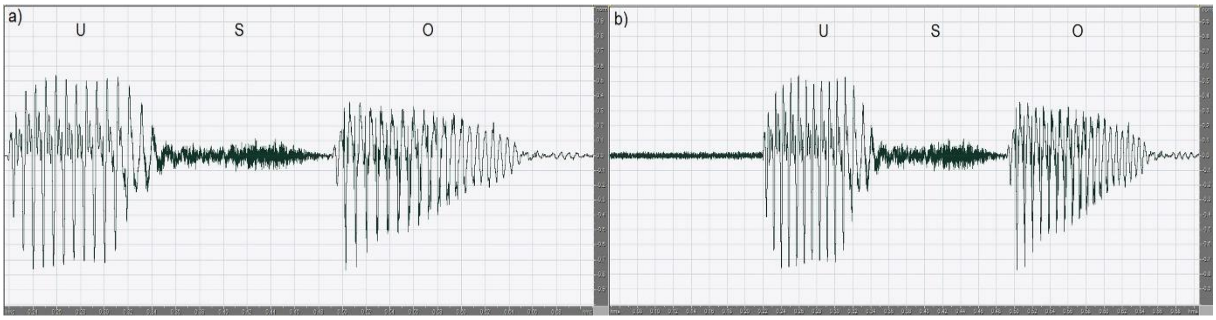

Fig. 1. The waveform of the logatom "Uso":
a) without the initial phrase, b)with the initial phrase.

Subjective measurements of speech signal quality are inseparably connected with the auditory perception process, which can be considered as the listener's response to the stimulus reaching him. The listener's response depends on this stimulus, as well as on the conditions in which the listener is. The listener's response from a physiological and psychological point of view can be considered in terms of a sensorial and emotional response. The sensory reaction is the result of a given stimulus exceeding certain thresholds of sensitivity or thresholds of hearing categories. A more complex and more difficult to analyze is the emotional response, which is not the result of the characteristics of the received signal, but the habits, individual features of the listener, as well as from his psychophysical state. One of the purposes of the use of logatomes was to eliminate the emotional reaction and leave only the sensory one - related only to the hearing mechanism.

2. Measurement of speech intelligibility by an alternative method with forced choice

Subjective speech quality measurements according to intelligibility criteria are based on sentences, words or logatoms lists. The results of subjective measurements of the quality of speech transmission should depend on the physical parameters of the tested transmission channel, and not on the structure of the language test. Logatom lists ensure the elimination of information at the semantic level and thus limit the impact of the listener's emotional response on the obtained result.

Measuring logatom intelligibility consists of sending nonsense-word lists, which are written down by the listeners. The correctness of saving is checked by an expert who calculates the average intelligibility of the logatom. The average logatom intelligibility is calculated as the ratio of the correctly received number of logatomes to all sent logatomes.

Subjective measurements of logatom intelligibility are recommended by the ISO recommendation: Acoustics -- The construction and calibration of speech intelligibility tests [11] and the Polish Standards: PN-T-05100 Analog communication systems. Requirements and methods for measurement of logatom intelligibility" [12] and PN-V-90001 Digital communication systems. Requirements and methods for measurement of logatom articulation [13].

The subjective measurement of logatom intelligibility by the traditional method is very time-consuming. In the Department of Acoustics and Multimedia at the Wrocław University of Technology, a new measurement procedure has been developed. This method is called a "modified intelligibility test with forced-choice" (MIT-FC) [14-16]. In the MIT-FC method, all experiments are controlled by a computer. The computer generates the selected (subsequent) logatom, and the listener selects the correct logatom from the list of alternative logatomes presented visually on the monitor screen. Based on experimental research, it was established that the number of seven displayed alternative logatomes allows to minimize the method error, while fully utilizing the listener's perceptual abilities [17]. Among the seven logatomes proposed to the listener is the presented logatom in a random position. The listener's task is to choose the correct answer. After the listener has made a selection, the program compares the selected logatom with the transmitted one. When the result of the comparison is positive, the value of the variable storing information about correctly received logatom is increased. After the measurements are completed, the result is displayed on the monitor screen.

In the MIT-FC method, logatoms are presented as isolated test units. Therefore, it is very important to choose the optimal duration of the sequence before the logatom appears.

3. Experiment

3.1. Subject

Ten people (5 women and 5 men) aged 23 to 28 with normal hearing participated in the experiment. All subjects were fluent speakers of the Polish language. All subjects participated voluntarily. An additional criterion for qualifying for research was good computer skills, especially fluency in typing on a computer keyboard.

3.2. Procedure

The speech was presented at a level of 60 dBA and was at least 20 dB above the threshold in quiet. Within an experiment, each logatom was presented only once. The participant of the study saved the heard logatoms on the computer keyboard. The response time was measured in the form of latency between the end of the logatom emission and pressing the first key. The research used a program written and launched as part of the diploma thesis [18]. The program made it possible to introduce interference and distortion with given parameters into the speech signal. Before starting the actual measurements, the participants had a two-hour training session.

Reaction time was measured in binaural listening conditions using high-quality headphones. One measurement session lasted 20 minutes, followed by a 10-minute relaxation break. The purpose of this break was to avoid the effect of fatigue on the results obtained. 100 logatome lists were used for the research. The interfering signal was the white noise of different levels (SNR -12 ÷ +40 dB). Tests to determine the reaction times were carried out for the following conditions:

- speech signal without noise and interference,
- noisy speech signal to varying degrees (SNR = 40, 30, 20, 10, 5, 0, - 2, - 4, - 6, - 8, - 10, - 12 dB),
- speech signal distorted by echo,
- speech signal noisy to varying degrees and distorted by the echo effect (delay T = 0.4, loss of amplitude decay A = 0.6)

All measurement conditions were carried out using the simulation method, using the program created as part of the thesis [18].

The choice of noise interference was determined based on previous studies on the analysis of the impact of noise interference on speech quality. An SNR value less than -12 dB gave logatome expression results below 10% [18]. In the case of echo distortions, 16 configurations of delay (T) - loss of amplitude decay (A) were examined.

4. Results

4.1. Reaction time for ideal conditions

The results obtained for the situation when the speech signal was neither distorted nor disturbed are presented in Table 1. The table uses the designation F – female, M – male; the number indicates the listener's number.

In the present study, it can be observed a significant difference in reaction time between male and female listeners. The average difference is about 177 ms, and the reaction time is much shorter for men than for females. Only in one case (female labelled F4), the response time is comparable to the response time of men.

4.2. Reaction time for different SNR values

The results obtained for the case when the speech signal was disturbed by the white noise of different levels are presented in Table 1 and graphically in Figure 2.

Tab. 1. Response times for ideal conditions.

	Female			Male		
Subject	Average reaction time [ms]	Standard deviation [ms]	Logatom intelligibility [%]	Average reaction time [ms]	Standard deviation [ms]	Logatom intelligibility [%]
1	745	120	89	398	108	94
2	643	83	94	377	62	93
3	768	102	88	491	234	88
4	468	156	92	600	243	92
5	631	160	91	505	147	96
Average value	651	124,2	90,8	474,2	158,8	92,6

Tab. 2. Response times for speech signal interfered with white noise.

	Female			Male		
SNR [dB}	Average reaction time [ms]	Standard deviation [ms]	Logatom intelligibility [%]	Average reaction time [ms]	Standard deviation [ms]	Logatom intelligibility [%]
40	659	232	90	474	206	90
30	689	257	85	483	174	86
20	708	201	87	494	179	85
10	759	256	72	535	215	76
5	820	253	60	558	220	60
0	835	246	47	575	196	56
-2	878	252	38	601	252	41
-4	937	317	36	665	308	39
-6	961	314	32	666	282	38
-8	978	307	23	679	291	26
-10	990	339	20	793	311	15
-12	1013	328	15	788	290	14

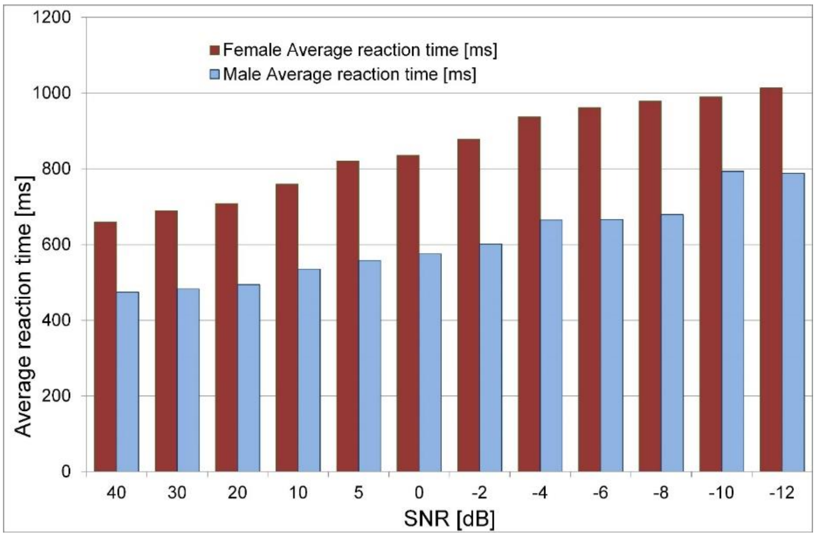

Fig. 2. Response time dependence on SNR for white noise interference.

This point of research showed that the response time in noiseless conditions is faster than the time obtained for SNR = 12 dB by 362 ms for female listeners and by 314 ms for the male group. The longest response time of 1 s was obtained in the female group for SNR = -12 dB.

4.3. Reaction time for echo distortion

The results obtained in the conditions of the speech signal with echo distortions for the combination of 4 delay times *T* (0.1s, 0.4s, 0.7s, 0.9s) and 4 values of amplitude decay *A* (0.1, 0.4, 0.7, 0.9) are given in Table 3 and graphically in Fig. 3.

Tab. 3. Response times for echo distortion.

T	*A*	Female			Male		
		Average reaction time [ms]	Standard deviation [ms]	Logatom intelligibility [%]	Average reaction time [ms]	Standard deviation [ms]	Logatom intelligibility [%]
T = 0,1	0.1	547	114	93	383	159	90
	0.4	567	106	89	424	172	96
	0.7	593	177	88	420	163	90
	0.9	600	195	84	442	164	90
T = 0,4	0.1	553	176	92	408	187	93
	0.4	587	181	92	431	163	90
	0.7	592	191	92	404	198	91
	0.9	619	185	88	435	156	90
T = 0,7	0.1	576	105	89	388	170	93
	0.4	599	160	90	444	163	90
	0.7	616	181	87	451	126	91
	0.9	619	117	79	482	148	89
T = 0,9	0.1	584	185	88	396	151	92
	0.4	599	174	93	450	146	93
	0.7	626	184	92	472	144	89
	0.9	633	136	79	488	115	83

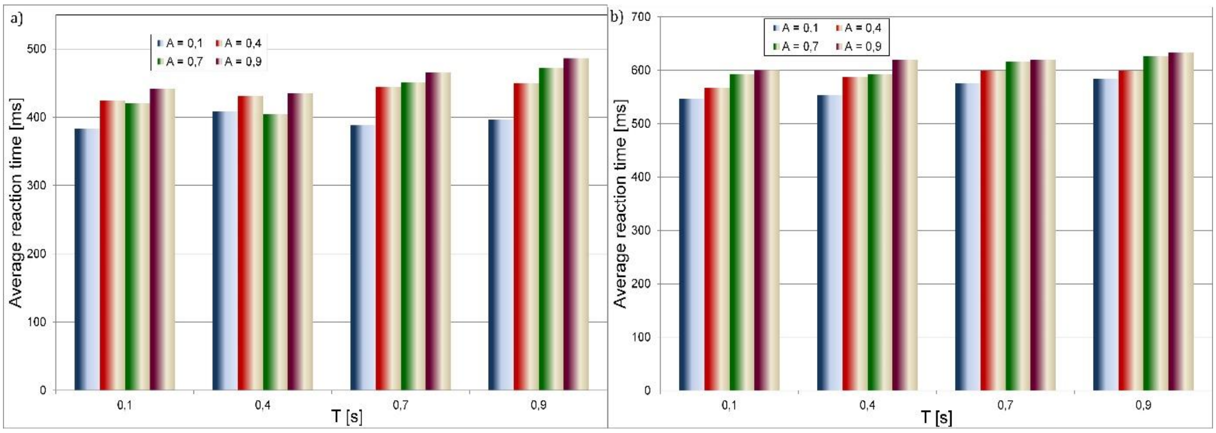

Fig. 3. The dependence of the response time in reverberation conditions as a function of the reverberation time for different amplitudes decay: a) male, b) female.

It was found that the reaction time increased slightly with increasing reverberation time. Using the ANOVA test at the significance level of α = 0.05, it was found that the reverberation time did not affect the reaction time in each group. Picou, Gordon and Ricketts [19] research also showed that the reverberation time has no significant effect on the response time.

4.4. Reaction time for various combinations of noise and distortion

In the part of the experiment concerning the influence of the reverberation time (T) and the amplitude decay (A), it was noticed that for the majority of listeners, especially in the male group, for the values of T = 0.4 and A ′= 0.7 there is a significant shortening of the response time to the value of amplitude decay A = 0.1 and 0.4. Hence, in the further part of the experiment, it was decided to set the reverberation time parameters T = 0.4, while the value of the amplitude decay was set to 0.6.

The results obtained under conditions of simultaneous speech signal echo for a delay time T = 0.4 s and decay of amplitude decay A = 0.6 and white noise of different levels are given in Table 4 and graphically in Fig. 4.

Tab. 4. Response times for speech signal distorted by echo and white noise interference.

SNR [dB}	Female			Male		
	Average reaction time [ms]	Standard deviation [ms]	Logatom intelligibility [%]	Average reaction time [ms]	Standard deviation [ms]	Logatom intelligibility [%]
40	639	204	91	427	192	91
30	642	206	89	462	192	92
20	659	227	86	464	184	91
10	708	235	76	482	232	79
5	748	248	62	563	251	69
0	753	264	56	620	273	54
-2	791	247	48	628	234	46
-4	856	259	40	634	223	39
-6	945	325	37	769	251	35
-8	938	287	26	760	281	26
-10	1020	355	18	792	314	24
-12	1085	366	10	832	304	15

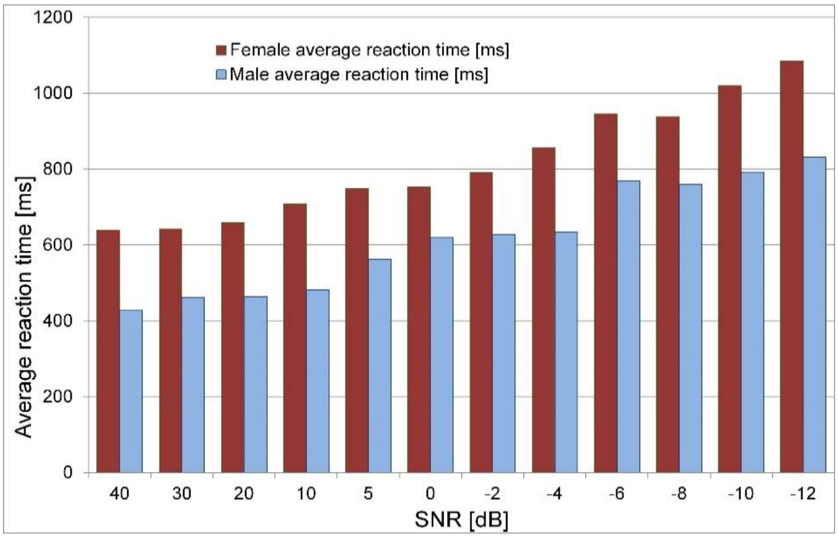

Fig. 4. Response time dependence on speech signal distorted by echo and white noise interference.

This experiment showed that background noise increases auditory effort, as evidenced by the longer response times in noise than in silence. A similar conclusion is drawn from the research carried out by Picou, Gordon and Ricketts [19]. In the present experiment, the shortest response time of 427 ms was obtained in the male group, while the longest 1085 ms was in the female group.

5. Conclusions

The main goal was to determine the reaction time of the listener, which would ensure the correct reception of logatomes. Considering that the group of students participating in the speech quality assessment consists of both females and men, when determining the length of the initial sequence of the logatom, the reaction time for the worst conditions should be taken into account. Analysis of the results shows that the time of the initial phase of the logatom should not be less than 1.1 s. Initial subjective speech quality tests carried out with the MIT-FC method showed that the introduction of the pre-sequence (noise with a level correlating with the noise level present in the recorded recording) significantly will improve the obtained results in terms of their compliance with the results obtained with the traditional method.

Within each test condition, the results of the obtained reaction time were subjected to statistical analysis. The homogeneity of the results obtained in the group of female and male listeners was checked using the Bartlett test, confirming the homogeneity at the significance level of α = 0.05, which allowed for averaging the ratings of all members of the listening team.

In the study of reaction time in the conditions of the noise of the speech signal with white noise, it was observed that if the first phone in the logatom has a noise character, then the listener's reaction took place only after the next phone, which distorted the latency time. This phenomenon was caused by the fact that noisy phones were not heard in conditions of a considerable high noise level.

During the study, an interesting phenomenon was observed for the perception of long logatomes, for which the response times were 60 ms, and even in some cases 30 ms. It turned out that some people participating in the experiment, after recognizing the first phone began writing logatom simultaneously analysing the rest of the transmitted logatom. It gave the impression that they heard the logatom before it was fully said. Because the response time was measured as the latency between the end of the logatom and the first keypress, the result was incredibly short response times. However, this phenomenon was very rare.

Additional information

The authors declare no competing financial interests.

References

1. Z. Borysiak, Types of stimuli (tactical, acoustic, visual) and sensomotor responses time in fencing, (in Polish: Rodzaje pobudzenia (taktylne, akustyczne, wzrokowe) a czas odpowiedzi czuciowo-ruchowych w szermierce), Medycyna sportowa, 1(6):28-34, 2006. http://31.186.81.235:8080/api/files/view/9807.pdf (21.05.2021).
2. W. A. Dreschler, Relations between psychophysical data and speech perception for hearing-impaired subject. II, Journal of the Acoustical Society of America, 78(4):1261-1270, 1985.
3. S. W. Link, Applying RT deadlines to discrimination reaction time, Psychonomic Science, 25:355-358, 1971 https://link.springer.com/content/pdf/10.3758/BF03335904.pdf (21.05.2021)
4. B. C. J. Moore, Introduction to the psychology of hearing,-Sixth Edition, Brill Academic Pub., 2013.
5. K. Schneider1, G. Dogil1, B. Möbius, Reaction time and decision difficulty in the perception of intonation, INTERSPEECH, Florence, Italy, 2221- 2224, 2011.
6. P. W. Stevenson, Reaction time measurements in speech discrimination tasks-an automated system with closed response sets, Journal of Phonetics, 1:347- 367, 1973.
7. N. J. Versfeld, W. Dreschler, The relationship between the intelligibility of time-compressed speech and speech in noise in young and elderly listeners, Journal of the Acoustical Society of America, 111 (401):401-408, 2002.
8. A. T. Welford, Reaction time, speed of performance and age, In: J. A. Joseph (Ed.), Central determinants of age-related declines in motor function, Annales of the New York Academy of Science, 515:1-17, 1988.

9. J. B. Gleason, N. B. Ratner, Psycholinguistics, 2 ed., 9 printing, Fort Worth : Harcourt Brace College, 2011.
10. S. Brachmanski, Chosen problems of speech transmission quality assessment (in Polish: Wybrane zagadnienia oceny jakości transmisji sygnału mowy), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, 2015.
11. ISO/TR 4870, Acoustics – The construction and calibration of speech intelligibility tests, 1991.
12. Polish Standard PN-90/T-05100, Analog telephone chains. Requirements and methods of measuring logatom articulation (in Polish), 1991.
13. Polish Standard PN-V-90001, Digital communication systems. Requirements and methods for measurement of logatom articulation (in Polish), 1999.
14. S. Brachmanski, The automation of subjective measurements of speech intelligibility in rooms, The 112th Conv. AES, Munich, Germany, 2002.
15. S. Brachmanski, Automation of subjective measurements of logatom intelligibility in classrooms, Automation/ed. by Florian Kongoli, InTech www.intechopen.com, 2012.
16. W. Majewski, W. Myślecki, K. Baściuk, S. Brachmański, Application of modified logatom intelligibility test in telecommunications, audiometry and room acoustics, Proceedings 9th Mediterranean Electrotechnical Conference Melecon'98, Tel-Aviv, Israel, 25-28, 1998.
17. S. Brachmański, Choosing the optimum number of test items in subjective logatom measurements (in Polish), Open Seminar on Acoustics OSA'95, Warszawa-Białowieża, 423-428, 1995.
18. M. Klinger, Influence of interference and noise to time of listener reaction, Masters thesis, Wroclaw University of Science and Technology, Wroclaw, 2008.
19. E. M. Picou, J. Gordon, T. A. Ricketts, The effects of noise and reverberation on listening effort for adults with normal hearing, Ear Hear, 37 (1):1-13, 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC4684471 (29.05.2021)